CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Upjohn Research

W.E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

Upjohn Research

Upjohn Institute Working Papers

Upjohn Research home page

1-28-2009

Estimating a Performance Standards Adjustment Model for Workforce Programs that Provides Timely Feedback and Uses Data from Only One State

Timothy J. Bartik
W.E. Upjohn Institute, bartik@upjohn.org

Randall W. Eberts
W.E. Upjohn Institute, eberts@upjohn.org

Kenneth J. Kline
W.E. Upjohn Institute

Upjohn Institute Working Paper No. 09-144

Follow this and additional works at: https://research.upjohn.org/up_workingpapers

Part of the Labor Economics Commons

Citation

Bartik, Timothy J., Randall Eberts, and Kenneth Kline. 2009. "Estimating a Performance Standards Adjustment Model for Workforce Programs that Provides Timely Feedback and Uses Data from Only One State." Upjohn Institute Working Paper No. 09-144. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. https://doi.org/10.17848/wp09-144

This title is brought to you by the Upjohn Institute. For more information, please contact repository@upjohn.org.

Estimating a Performance Standards Adjustment Model for Workforce Programs that Provides Timely Feedback and Uses Data from Only One State

Timothy J. Bartik
Randall Eberts
Ken Kline

July 2004
(revised January 28, 2009)

The W.E. Upjohn Institute for Employment Research
300 South Westnedge Avenue
Kalamazoo, MI 49007
contact: bartik@upjohninst.org

Authors' Note: This paper formed the basis for a performance reporting system that was developed for the State of Michigan. The system is referred to as the "Value-Added Performance Improvement System" (VAPIS) and has been implemented for more than a year. The system adjusts the U.S. Department of Labor's common measures for WIA workforce programs for factors that are beyond the control of local administrators, such as the characteristics of program participants and local labor market conditions. The common measures include three labor market outcomes: entered employment, job retention, and earnings levels. By making these adjustments, the common measures more closely approximate the value added that the workforce programs contribute to the labor market outcomes of participants. VAPIS also provides a short-term forecasting component that assists local workforce administrators in understanding the likelihood of that their current participants will find and retain jobs. Because of the long lag in reporting common measures, local administrators have little systematic knowledge of their performance. VAPIS tries to fill that gap.

The purpose of this paper is to describe a methodology for adjusting performance standards for workforce programs offered by local workforce areas (LWAs). By performance standards adjustment, we mean a model that uses a statistical approach to attempt to better measure the relative performance of different local workforce areas in providing workforce system customers with "value added" in terms of the system's desired outcomes. Our paper's approach has four distinguishing features. First, the performance standards are based on the common measures proposed by the U.S. Department of Labor, which include short- and longer-term employment outcomes. Second, the model is estimated using data from only one state, which allows each state greater flexibility in adapting the adjustment model to the state's needs and available data. Third, the model is estimated using data on individual customers, which offers some estimation advantages, particularly when data from only one state is available. Fourth, since some of the common measures are not available until long after the program year is completed, we include real-time predictions of the current performance of the LWA and an assessment of whether or not it will meet its performance standards when the common measure data is eventually available. This more timely feedback on performance provides administrators the opportunity to better manage their operations and offer services that best meet the needs of their customers.

Under the Workforce Investment Act of 1998, performance standards for state workforce programs and local workforce areas have been based on negotiations between the various parties. Each state negotiates with the U.S. Department of Labor to set standards for each of several performance measures for the program year. The states in turn negotiate with each of their LWAs to determine their performance standards for the same period. As this practice of setting standards evolved, states and LWAs increasingly believed that this method was not taking into account factors that affected their performance but were beyond their control, such as the conditions of the local labor market, or factors that could be manipulated to improve performance but did not reflect the true effectiveness of the services provided, such as the characteristics of the customers enrolled by the LWAs.

During the time this practice was receiving increased scrutiny, the Office of Management and Budget embarked on an initiative to improve the management and performance of the Federal government. Part of this initiative included the implementation of common performance measures across the Federal job training and employment programs. The purpose of proposing a more comprehensive and integrated system was to enhance the ability to assess the effectiveness and impact of the workforce investment system.

Seeing opportunities to develop strategies and guidance for state/local workforce investment system goal setting and performance adjustment, the State of Michigan received a grant from the Performance and Results Office of the U.S. Department of Labor, Employment and Training Administration. One component of this grant was to develop a statistical methodology to provide a more accurate assessment of the true effectiveness, or "value-added," of the workforce system. The model focuses on the common measures proposed by the Office of Management and Budget and endorsed by the U.S. Department of Labor (TEGL No. 15-03, December 10, 2003). The common measures are based on the employment outcomes of participants of the programs offered by an LWA. However, these measures cannot be used to assess the effectiveness and impact of each local office without appropriate adjustments since they also include the effects of local labor market conditions and the abilities and qualifications of the participants themselves. Therefore, using unadjusted performance measures entangles the true value added effects with the effects of local labor markets and personal characteristics. As pointed out by some states and LWAs, it is conceivable that an LWA may be credited with greater success than others, not because its services are more effective, but simply because the individuals it enrolls are better qualified to obtain or retain a job or because the local economy is more robust in creating jobs. Therefore, to fairly judge the performance of LWAs, the true value added of the services they provide must be separated from the other factors.

To do this, one can ask the hypothetical question: How would the performance of two LWAs compare if they provided services to participants with identical characteristics and subject to the same labor market conditions? By making such a comparison, the only difference in performance would be in the effectiveness of the delivery of services. While making pairwise comparisons is possible, it is not feasible on a large-scale basis. Furthermore, comparisons do not provide information on how the various factors affect performance, which can help state and local administrators better understand the reasons for their performance. A more practical approach is to use statistical methods to adjust the performance of LWAs so that they can be evaluated according to the true value added of their services, separate from the other factors.

Statistical adjustment to performance standards is not a new concept. It was used under JTPA to adjust the performance standards for each Service Delivery Area (SDA). Key factors reflecting the characteristics of participants and the local labor market conditions were included in the JTPA adjustment model, which was estimated at the national level using data from each SDA on the SDA's performance and the SDA's means for participant characteristics and local labor market characteristics. The coefficients obtained from the national model were then used to weight the value for each factor at the SDA level.

The adjustment methodology proposed in this paper has some similarities to the JTPA approach. A major difference is that instead of using data aggregated to the SDA level and estimating a model for the nation, we use data on each participant and estimate a model for a single state. One reason for the different approach here is that the common measures are based on Unemployment Insurance wage records and these records are not available at the national level but only on a state-by-state basis. There are also other reasons why it may be desirable to have a different adjustment model for each state. States may differ in the way in which personal characteristics and local labor market conditions affect performance outcomes, which implies that each state should have a different model. In addition, allowing each state to develop its own adjustment model gives each state more flexibility in setting up its performance standards system.

If an adjustment model is estimated using data from only one state, it is infeasible to estimate the model using data on LWA means. Because the number of LWAs that exist in one state is limited, estimating a model using data on only LWA group means would result in estimates that would be too imprecise. Therefore, this paper estimates an adjustment model using data on individual participants.

A good adjustment model for workforce performance standards should lead to a better measure of value-added of an LWA. In turn, such an approach will allow the resulting performance standards system to better promote higher value-added among LWAs by better identifying high-value-added LWAs that should be rewarded and emulated, and low-value-added LWAs that should be reformed. By adjusting for personal characteristics, a good adjustment model avoids encouraging LWAs to "cream-skim" the local population, seeking to enroll only those workforce participants who would have done well anyway. Finally, at least before the results of the system are known, most LWAs and the public are likely to agree that a well-designed adjustment system is fairer, because it holds LWAs harmless for trying to serve those who need greater assistance and thus are less likely to find or retain a job. In so doing, the need for LWAs to meet standards is more compatible with trying to serve the needs of customers the best way possible.

In order to capture the "longer-term" effects of workforce programs, the common measures include employment and earnings information for up to three quarters after the participant exits a program. While the wage record data have advantages, such as being comparable across LWAs and states and containing fewer non-response problems than would be true of individual household surveys, the data are available only with considerable lag, perhaps two to three quarters after the quarter being examined. This delay is problematic for a performance standards system because the feedback of an LWA's performance might be a year and a half after the program year is over, by which time the program and the LWA might already have changed too much for the feedback to be relevant. To deal with this problem, this paper also includes a forecasting model that predicts performance on the common measures using data on program participants that are available by the time they exit a program. These forecasts, combined with adjustments based on the adjustment model, can be used to provide each LWA administrator with quick "real-time" predictions, during and shortly after the program year, for how the LWA will eventually do on the adjusted performance measures. These quick "real time" predictions are likely to encourage LWA administrators to implement more relevant and effective program modifications.

The results presented in this paper, which are estimated for a variety of workforce programs in Michigan, suggest that many individual participant characteristics and local economic conditions do help to predict how well individual workforce system participants will do on the common measures. The average unadjusted common performance measures of LWAs are often highly correlated with predictions based on these estimated effects of participant characteristics and local economic conditions, which suggests that adjustment of performance is necessary for an accurate performance standards system. However, the empirical work also suggests that even after adjustment, there is still considerable true differences in "value added" across LWAs, which rationalizes having a performance standards system to identify high-value-added and low-value-added LWAs. Finally, for at least some of the common measures and groups, it is feasible to provide reasonably accurate real time predictions of an LWA's value added. Suggestions for how to improve such predictions will also be considered in this paper.

The Models

The common measures analyzed in this study are defined in Table 1. Since our model is based on individual data and not LWA-aggregated data, some practical modifications had to be made to the definitions. As described later, these modifications do not affect the measured performance of LWAs.

Using individual data, we seek to estimate how much of the variation in individual performance on the common measures is due to individual characteristics and local economic conditions and how much is due to the "value added" of the LWA. The model can be written as:

(1) $$Y_{ij} = \boldsymbol{B}\boldsymbol{X}_{ij} + W_j + e_{ij}$$

The model is estimated by pooling data across individuals (indexed by i) participating in programs offered by different LWAs (indexed by j). A separate model is estimated for each common measure and program. Y_{ij} denotes one of the common measures for individual i in LWA j. $\boldsymbol{B}$ is a vector of coefficients to be estimated. A different set of coefficients would of course be estimated for each common measure and for each program. $\boldsymbol{X}_{ij}$ denotes a set of relevant individual characteristics that might affect performance on the common measures, such as an individual's age, education, prior earnings history, gender, or race. In addition, the $\boldsymbol{X}$ variables include the change in unemployment that occurred in that local area over the relevant time period covered by the common measure for that individual. As can be seen in Table 1, all of the common measures implicitly refer to some change in employment or earnings or educational attainment for an individual between two time periods.

It should be recognized that $\boldsymbol{X}_{ij}$ would not reflect events that happened to the individual after they entered the program, such as what services they were assigned to or received, how well they did in that service, and any intermediate outcomes measuring success, as these are assumed to be part of the relative value added of the programs that we want to estimate separately.

W_j represents a fixed effect for LWA j. The fixed effects for each of the LWAs are estimated by including a complete set of dummies for each of the LWAs in the estimating equation. Ideally, if it were possible to measure all relevant individual characteristics and local economic conditions that affect individual performance on the common measures, the estimated W_j would be an unbiased and consistent estimate of the relative value added of the LWA in terms of how the LWA's performance affected the common measures. Of course, because there are omitted individual characteristics or local economic conditions that may not be fully reflected by the measured $\boldsymbol{X}$ variables, and these omitted variables may differ across LWAs, the estimates of the W_j effects may be subject to some bias. These statistical issues are considered further below. But we want to emphasize at the outset that even if there is some bias in the estimates of the W_j effects, the relevant issue is whether these estimates are still closer to the true relative value added than the estimates one would obtain by simply comparing LWA means on the common measures, which is equivalent to estimating Equation (1) with all $\boldsymbol{X}$ variables omitted.

If we had data on individuals who did not participate in any workforce program (in essence a control group), and if we could be assured that on average these individuals did not differ from program participants with respect to the omitted variables, then the W_j effects could be normalized by setting the W_j for this control group to zero, and the W_j for each LWA would actually measure the value added of each LWA program relative to the outcomes of not participating in the program. But for the models estimated in this paper, we only have and use data on program participants. This is partly a matter of data convenience, as it is costly and difficult to obtain administrative data on individual characteristics for non-program participants. In addition, there is a judgment here that any control group of non-participants, absent the creation of a random assignment experiment, is likely to differ in many omitted variables from program participants. In contrast, differences across LWAs in characteristics of program participants are more limited by the design of the program.

To estimate the W_j effects, we need some arbitrary normalization of the W_j effects. To do so, we constrain the weighted sum of the W_j effects to be zero, where each W_j is weighted by the number of participants in that program for that LWA. This normalization means that each estimated W_j implicitly seeks to measure the value added for that LWA, for that common measure and program, relative to the weighted state average value added, for that same common measure and program. The equation includes a disturbance term, denoted by e_{ij}.

Ex post, after some particular program period is completed, program data can be used to estimate a model such as Equation (1). This provides ex-post estimates of the relative value added for each LWA, compared to the weighted state average. Some performance standard can be defined relative to that state average, and it can be estimated whether each LWA has met that performance standard. So, for example, we could say that an LWA had only met the performance standard for that common measure and program if the relevant estimated W_j was greater than or equal to some cutoff M. If M is zero, then the implicit standard is state average value added. If M is negative, the standard is some value-added level that is below the average state value-added, and if M is positive, the standard is some value-added level that is above the average state value-added. How exactly to set M is a difficult issue. Presumably, one would want to set M so that most but not all LWAs are able to meet the performance standard, but then gradually ratchet that standard up over time.

Of course, this ex-post approach would require waiting until after the program year when all the common measure data were available for analysis, and then waiting around beyond that time while some group of researchers used these data to estimate the adjustment model. We want to have adjustment results and estimates of value-added sooner than that. To see how to do this, it is helpful to first consider some other equations that follow from the model presented in Equation (1), and then to use these equations to develop other ways to state the performance standard.

The estimates of Equation (1) will always have the following two features:

(2) mean $Y_j = \boldsymbol{B}$ (mean $\boldsymbol{X}_j$) + W_j

(3) mean $Y_s = \mathbf{B}$ (mean $\boldsymbol{X}_s$).

Equation (2) says that the mean of Y for a given LWA indexed by j is equal to the estimated $\boldsymbol{B}$ times the mean of $\boldsymbol{X}$ for that LWA, plus the estimated W_j for that LWA. Equation (3) says that the mean of Y for the state is equal to the estimated $\boldsymbol{B}$ times the mean of $\boldsymbol{X}$ for the state.

Using these two equations, setting a standard that W_j is greater than or equal to M, where M is some value added standard relative to the state average value added, is equivalent to saying the standard is:

(4) mean $Y_j \geq \boldsymbol{B}$ (mean $\boldsymbol{X}_j$)+ M or

(5) mean $Y_j - \boldsymbol{B}$ (mean $\boldsymbol{X}_j$ − mean $\boldsymbol{X}_s$) ≥ mean $Y_s + M$.

Inequality (4) says that the performance standard can be restated as requiring that an LWA's mean performance must exceed what we would predict for that LWA based on its mean $\boldsymbol{X}$s, plus the M modification to state average performance. Inequality (5) says that another way to restate that performance standard is to say that the LWA's mean actual performance on the common measure Y, when adjusted by an adjustment factor (the second term), must be greater than or equal to the state mean, when modified by M. For the kth variable X_{ijk} that has a positive effect on the common measure, and hence a corresponding positive estimated coefficient B_k, if the LWA is better than the state average (mean X_{jk} − mean X_{sk} is greater than zero), then this adjustment subtracts some number from the LWA's performance before comparing it with the standard of mean $Y_s + M$. This makes sense because if an LWA is better than the state average in having participants who are expected to do better on the common measure even without the program, or if an LWA experiences a better-than-average economy, one would want to adjust downwards the LWA's actual performance before comparing it with a fixed standard. On the other hand, if for some variable that positively affects the common measure, this LWA is worse off than the state average, then this adjustment will positively increase the area's score above its actual mean before comparing it with a fixed performance standard. Another way to write the adjustment factor in Inequality (5) is as [$\boldsymbol{B}$ (mean $\boldsymbol{X}_j$) − mean Y_s], which is sometimes more convenient.

So far, we have written these inequalities as if there is some fixed statewide standard performance, and we take each LWA's actual mean performance and adjust it based on its $\boldsymbol{X}$s to see if it meets that fixed standard. An exactly equivalent method to determine whether each LWA meets its performance standard can be derived that instead compares the actual mean performance of each LWA with a performance standard that is adjusted based on each LWA's $\boldsymbol{X}$s. This is done by shifting the adjustment factor in Inequality (5) to the other side of the inequality, which says that the performance standard is met if the following inequality holds:

(6) $\text{mean } Y_j \geq \text{mean } Y_s + M + \boldsymbol{B}\,(\text{mean } \boldsymbol{X}_j - \text{mean } \boldsymbol{X}_s)$

The intuitive interpretation of Inequality (6) is that if an area is better off on characteristics that positively affect the common measure, we should adjust its performance standard upward. If the area is worse off, we should adjust its performance standard downward.

Suppose that instead of telling whether the LWA meets a performance standard, we simply want alternative ways of expressing W_j, the area's value-added relative to the state average. Then the following two equations are useful re-expressions:

(7) $W_j = \text{mean } Y_j - \boldsymbol{B} \text{ mean } \boldsymbol{X}_j$

(8) $W_j = (\text{mean } Y_j - \text{mean } Y_s) - \boldsymbol{B}\,(\text{mean } \boldsymbol{X}_j - \text{mean } \boldsymbol{X}_s)$

Equation (7) is intuitively interpreted as saying that relative value added is the extent to which the LWA's mean performance exceeds what would have been expected based on the area's mean characteristics. Equation (8) is interpreted as saying that an LWA's relative value added compared to the state average is equal to its performance relative to the state average, minus the already-used adjustment factor that reflects the area's characteristics relative to average state characteristics.

So far, although these equations and inequalities allow for some interesting interpretations of the meaning of relative value-added and adjusted performance, all of this analysis just shows alternative ways of calculating the same concepts, using ex-post analysis and estimation with program and wage record data. But if we want to estimate relative value-added and adjusted performance measures in a more timely manner, we can instead use the estimated $\boldsymbol{B}$ and statewide means of $\boldsymbol{X}$ and Y from an earlier historical period, which we assume have already been estimated by some researchers and are available prior to the program year being analyzed. Using these historical data, which we will designate using a b ("before") subscript, an estimate of relative LWA value added is given by:

(9) Estimated W_j = (mean Y_j − mean Y_{sb}) − $\boldsymbol{B}_b$ (mean $\boldsymbol{X}_j$ − mean $\boldsymbol{X}_{sb}$)

Furthermore, one estimate of whether the LWA met a relative performance standard is to require that its mean performance satisfy the following inequality:

(10) mean Y_j − $\boldsymbol{B}_b$ (mean $\boldsymbol{X}_j$ − mean $\boldsymbol{X}_{sb}$) ≥ mean $Y_{sb} + M$

In using the statewide means of Y and $\boldsymbol{X}$ from a "before" or historical period, these calculations are essentially measuring an LWA's relative value added compared to average state value-added during this historical period. There is nothing wrong with such a comparison. There are some advantages in basing performance standards on data that one has had some time to analyze and decide what are relevant standards to set. It seems preferable to base the levels of standards on historical data, which allows all LWAs in the current year to have a chance to meet the performance standards. Basing performance standards completely on current year data essentially imposes a "curved" grading system, in which inevitably some LWAs must fail the performance standard.

In general, the $\boldsymbol{B}_b$ coefficients, estimated using the historical data, will differ from the $\boldsymbol{B}$ coefficients that would be estimated using the current program year's data, but that does not mean that one set of coefficients is necessarily superior to the other. In an ideal world, in which these $\boldsymbol{B}$ coefficients truly capture all relevant control variables affecting performance, the W_j coefficients are truly unbiased estimates of value-added, and sample sizes for all programs are large enough that the imprecision of estimates can be ignored, there may be some argument for using the current year $\boldsymbol{B}$ (and implicitly the W coefficients), as the structure of the model may change over time. However, in the real world, with many omitted variables potentially biasing the estimates, it is unclear whether the historical or current estimates are better. Furthermore, using historical data should in general allow pooling several program years, which should allow more precise estimates than just using the current program year.

So far, we've avoided the time-consuming step of getting some researchers to estimate the B coefficients using current-year data, but we still have to get current year data on Y and $\boldsymbol{X}$ for each LWA to estimate value-added and to determine whether the performance standard has been met. The complete set of actual Y and $\boldsymbol{X}$ data will of course be available only with some lag. But it may be possible to predict likely values of $\boldsymbol{X}$ and Y much sooner.

The $\boldsymbol{X}$ variables are easier to predict early in the program year. In general, all of the individual characteristics included in the $\boldsymbol{X}$ variable list, and hence used for adjustment, should probably be characteristics observed at or close to registration. We certainly do not want to control for any individual characteristics that represent effects of the program rather than inputs. For example, in making adjustments we would not want to control for the services the participant received and whether they did well in the training modules. Such individual characteristics represent part of what the program is doing that constitute its value-added, and would not be included in the ***X*** variable list. Therefore, at registration or shortly thereafter we should actually observe all the individual characteristic ***X*** variables, and therefore be able to calculate a running "real time" value of the mean ***X*** in each LWA.

The only ***X*** variables that will not be known with certainty near registration are ***X*** variables that represent the effect of the economy. In any reasonable model of what local economic factors affect common measures that reflect labor market outcomes in the first quarter or third quarter after exit, one would have to allow for local economic characteristics as of the first or third quarter after exit. Obviously these local economic characteristics will not be observed until some time after the first or third quarter after exit, given the time necessary for economic statistics to be collected and verified. However, it is certainly possible to forecast what will happen in the local economy. These forecasts of local economic conditions can be plugged into Equation (9) and Inequality (10) to provide an early estimate of the "adjustment factor." Assuming the influence of local economic characteristics does not loom "too large" compared to the influence of individual characteristics in affecting the common measure, and assuming the forecast error is not "too large," these early estimates of the adjustment factor will be reasonably close to their final actual value.

For the present empirical illustration of this model, the only local economic condition variable that is included in the model is the change in the local unemployment rate over the time period encompassed by the common measure. As will be discussed below, in doing illustrative calculations we assume that the forecast change in unemployment is zero. It would be possible to potentially improve the estimates by developing a better forecast of the change in the unemployment rate.

The actual values of the average common measures for each LWA, the mean (Y_j), will not be known for sometime after the program year, as the wage record data used is from one to three quarters after exit, and wage record data are only available with some lag. But as with the $\boldsymbol{X}$ variables, an early forecast of Y can be produced during or shortly after the program year, using data available administratively at that time. For example, it would seem reasonable to forecast the Y variables using administrative data on various "intermediate outcomes" for program participants. The most obvious intermediate outcome to use in forecasting is whether the individual is employed at exit from the program, which seems a plausible predictor of whether the individual will be employed one quarter after exit, which is the "job entry" common measure. Workforce programs also frequently have data on the wage rate, hours of work, and occupation of the job the participant held at exit. We would presume that such variables would help predict the participant's earnings gains and job retention, which are other common measures.

Using historical data on the program, we would estimate equations of the following form:

(11) $Y_{ij} = \boldsymbol{C}\boldsymbol{Z}_{ij} + V_j + u_{ij}$

The $\boldsymbol{Z}_{ij}$ variables include all the $\boldsymbol{X}$ variables, plus intermediate outcome variables that were not included in $\boldsymbol{X}$ because they might be an effect of program "value added." The V coefficients represent unmeasured fixed effects of each LWA that are not captured by the $\boldsymbol{Z}$ variables, but the V coefficients should not be interpreted as a measure of program value added. With estimates of $\boldsymbol{C}$ from this historical period, and measures of $\boldsymbol{Z}$ as each individual goes through the program, the expected mean value of Y can be forecasted for each LWA by either of the following two equations:

(12) forecast mean $(Y_j) = \boldsymbol{C}_b$ mean $(\boldsymbol{Z}_j)$

(13) forecast mean (Y_j) = $\boldsymbol{C}_b$ mean ($\boldsymbol{Z}_j$) + V_{jb}

The final mean of $\boldsymbol{Z}$ for the LWA won't be known until the end of the program year, but a running mean or real-time prediction can be calculated as individuals exit the program throughout the program year. The two forecasts differ in assumptions about whether the previous unobserved fixed LWA effect, controlling for observed intermediate outcomes, is more likely to persist unaltered (Equation 13) or revert to the state mean (Equation 12). The current empirical illustration of the model assumes the latter.

These forecasts of Y would then be adjusted by forecasts of the adjustment factor, as outlined before, to allow estimates of Equation (9) and Inequality (10), the program's value added and whether the program will meet the performance standard. Predicted value added and performance relative to standard can be estimated on a real-time basis throughout the year, as additional individuals enter and exit the program.

Estimation Details

There are many important details in how these models were implemented in the current empirical illustration, using workforce data from the state of Michigan.

Estimation method. For simplicity and speed and because of the large numbers of models estimated, all these adjustment models and prediction models are estimated using linear regression models, even when the dependent variables is a zero-sum variable. Using logit or probit would make it more difficult to interpret results and creates some complexities in calculating adjustments. For example, because logit and probit are non-linear models, the adjustment factor cannot be calculated using sample means for the LWA and the state, but rather requires calculating probabilities for all observations using the full set of data. Some experimentation indicates that in practice the more complicated logit or probit approach makes little difference in the resulting estimates of value-added.

Choosing the ***X*** and ***Z*** variables to use in the adjustment and forecasting models. In estimating all models, we aggressively reduced the number of variables in both the adjustment models (Equation 1) and the forecasting models (Equation 11). Variables with t-statistics less than 1.4 in absolute value were excluded from the model. This excludes variables that on an individual basis would be excluded from the model based on the Akaike Information Criterion, which is a model selection criterion that seeks to maximize the ability of a model to do out-of-sample predictions (Amemiya 1985: 44–55). As both the adjustment and forecasting model would be intended to do out-of-sample predictions, such model selection seems appropriate. In addition, for the adjustment model, we dropped variables with the "wrong" sign, even if statistically significant. The theoretical rationale is that the "wrong sign" indicates that the variable's coefficient is somehow being biased by some omitted variable bias. The practical rationale is that it is undesirable from a political or just good public policy perspective to make adjustments to calculate value-added or performance standards that go in the wrong direction. We do not want to penalize LWAs that serve more disadvantaged individuals, even if some estimated model might indicate that these more disadvantaged individuals do better on some common measure. We prefer to drop this variable and interpret this as an improvement in model specification that also yields results that are more relevant to day-to-day program use.

Calculating Percentage Change. The original common measures, as outlined by ETA pursuant to OMB guidelines (USDOL, ETA: TEGL, December 10, 2003), clearly envision using grouped data to calculate the common measures that are based on "percentage change" earnings gains. But the adjustment and forecasting models used here rely on individual data. Therefore, percentage change in earnings must be calculated for each individual. A conventional percentage change calculation results in infinite values in some cases or unreasonably large positive or negative values for some observations. For example, for common measure 3, which measures percentage change from the first quarter before registration to the first quarter after exit, and uses the first quarter before registration as a base, the percentage change is infinite for anyone who had zero earnings in the first quarter before registration. Even if one switches to an average base from the two time periods, the percentage change, when applied to an individual, results in putting very great weight on percentage changes that represent very small change in dollars for the individual. For example, under the percentage change approach using an average base to calculate percentages, a change in earnings from $200 to $500 per quarter is the same percentage change as a change from $2,000 to $5,000 per quarter. For policy purposes the latter change is far more significant.

To avoid this problem, we calculate the actual change in earnings between the two periods for the individual, and convert to percentage change for each observation by using the state mean for program participants during the initial period in the denominator. The mean of this variable for the state will equal the mean percentage change using state means in the two time periods. The implications of this for calculating value-added and performance standards is that implementation of this adjustment model and forecasting model requires that the U.S. Department of Labor be willing to accept these definitions of percentage change as within the intent of the common measure guidelines. In particular, to implement these adjustments and forecasts on a real-time basis, the state means from the historical period would have to be used as the base in calculating percentage change for the current program year.

Approximating the youth common measures. For the two youth common measures that we addressed, current Michigan administrative data do not allow these variables to be perfectly captured, only approximated. (For the literacy and numeracy gains measure, youth common measure 3, no approximation is even possible as there currently is no follow-up data on literacy and numeracy gains.) For the youth common measure 1, “entered employment or advanced education/training,” this variable is supposed to reflect whether the individual was employed or in the military, education, or advanced training as of one quarter after exit. However, no Michigan administrative data are available on involvement in the military, education, or advanced training as of one quarter after exit. Therefore, for purposes of the analysis in this paper, an individual was deemed to be a “one” on this variable if he or she were either employed in the first quarter after exit according to wage record data, or if his or her “exit reason” was that they entered the military, or entered apprenticeship, advanced training, or post-secondary education. For the youth common measure 2, “attainment of educational/training credential,” this variable is supposed to reflect whether the individual attained a diploma, GED, or certificate by the **end of the third quarter after exit**. However, Michigan administrative data do not currently report such a long-term follow-up. Rather, this variable is defined based on whether an individual had attained a diploma, GED, associates degree, bachelor or master’s degree, occupational certificate, or occupational license, as of **program exit**.

ES Common measure definitions. For training programs, it makes sense to define common measures relative to exit, as by definition the program goal is not to achieve immediate employment, but only after the individual has received training allowing them to have a better job or career. But for the employment service, the policy goal seems to be more to achieve immediate employment. This implies that ES common measures might more appropriately be defined relative to the individual's registration. We did all analysis of the ES both ways, defining the common measures both relative to exit and relative to registration. At least for Michigan data, the results showed little difference, so we have reported only the ES exit-based results in this paper. However, we think that policymakers might consider whether the common measures require some modification for the ES.

Correcting for uncertainty about who will be in the final sample for the common measures. One problem in using the adjustment model and forecasting model to do real-time predictions of adjustments and value added, during and shortly after the program year, is that we don't know fully at registration, or even at exit, who exactly will be in the sample that counts for calculating performance standards. Individuals are excluded from any of the common measures if they exit for certain reasons, for example if they are imprisoned or hospitalized (see note to table 1 for complete list). For adult common measures 2, 3, and 4, individuals are included only if they have positive earnings one quarter after exit.

The simplest alternative to dealing with this problem is to ignore it, by calculating the adjustments and the forecasts used for these real-time calculations using all individuals in the program. The hope would be that the later exclusion of some of these individuals from the final sample will not bias the real-time calculations too much.

However, we chose a somewhat more complicated alternative to do this adjustment, which is to estimate for each individual what their probability is of being in the final sample. To do real-time calculations of the adjustment factor, we estimated for each common measure a logit equation to predict whether they would be in the final sample for that common measure. For example, for adult common measures 2, 3, and 4, this logit model is similar to what was estimated for common measure 1, because the main reason people are excluded from common measures 2, 3, and 4 is that they were not employed one quarter after exit. The model is not identical to the common measure 1 model because the common measure 2 through 4 sample includes individuals employed at registration, whereas the common measure 1 sample does not. Once this logit model is estimated, it is used to weight each observation in calculating the LWA means and predicted values, with the weight equal to the probability that the observation is in the final sample. At exit, a new logit model to do this weighting is again estimated, because at exit we have more information about whether the individual is likely to be in the final sample; for example, for adult common measures 2, 3, and 4, at exit we know whether they are employed, which helps predict whether the individual will be employed one quarter after exit. So at exit, new probability weights are estimated and used to calculate a new adjustment factor and a new weighted forecast of the common measures. In all these weighing exercises, logit is used, rather than linear regressions, because it is important for calculating weighted means or weighted predictions to have weights that can never be negative.

In-sample versus out-of-sample implementation of the model. As outlined above, ultimately these adjustment models and forecasting models are designed to be estimated on program data for a historical period, and then used to predict adjustment factors, common measure outcomes, and value added for an out-of sample period. For the present paper, these predictions were done using the same data used to obtain the original program estimates. In other words, all parameters were estimated using all the program data available to us. We then predicted the adjustment factors and the common measure outcomes and value added on a real-time basis by "pretending" that we did not know the change in unemployment for each individual (in fact, we assumed it to be zero), and that we did not know who would be in the final sample. Using the same sample for both estimation and testing probably exaggerates the forecasting capabilities of the model for out-of-sample forecasts. On the other hand, this approach allowed us to use a much larger sample for parameter estimation than would otherwise be possible, allowing for more accurate estimates.

Critiques of the Model

We consider three possible criticisms of the model.

<u>An adjustment model leads to standards that are "moving targets."</u> One possible criticism is that the adjustment model approach leads to standards that will not be known prior to the program year, and in fact will change over the program year. But this "problem" is inherent in any performance standards system that is responsive to the characteristics of the individuals who actually enroll in the program and the actual labor market conditions experienced by the LWA. If policymakers want to design the performance standards system so that it recognizes the difficulties posed by a sudden influx of more-difficult-to-serve program participants, or a sudden downturn in the local economy, they must have a system that is adjusted during the program year.

From another perspective, a performance standards adjustment system that adjusts to the average characteristics of the LWA's program participants and local economy is attempting to always measure as accurately as possible the true value added of the LWA. A performance adjustment system that accurately measures the true "value added" of an LWA for a program will in some sense keep the same standard and same target throughout the program year. Even as the "nominal" standard is adjusted to respond to changes in participant mix and local economic conditions, the difficulty of the "real" standard, which is whether the true value added exceeds some cutoff, will stay the same. In practice, these adjustments will be imperfect, but a good adjustment system will keep the real standard more stable than a system with no adjustments.

Omitted variables may bias measures of value added and adjustments to performance standards. As in any model, some variables will inevitably be omitted from the adjustment model, which may bias the model's estimates of value added and performance standard adjustments. Here, the omitted variables would include some characteristics of the program participants, and some features of the local economy. However, a standard analysis of bias due to omitted variables suggests that this omitted variable bias is only important to the degree to which there are differences across LWAs in omitted variables that are not reflected in the included variables. Even if there is some significant bias in estimates of value-added and performance standards adjustment, the resulting estimates are likely to be better than the implicit estimates when no adjustment model is used.

Suppose the true model that we should be estimating is given by the following equation:

(14) $Y_{ij} = \boldsymbol{BX}_{ij} + \boldsymbol{DS}_{ij} + W_j + e_{ij}$,

which is identical to the original adjustment model Equation (1), except that now we are assuming that some "omitted variables" $\boldsymbol{S}_{ij}$ should be included. Then, as shown by a standard analysis of omitted variable bias, omitting $\boldsymbol{S}$ when it should be included yields the "biased" estimators of $\boldsymbol{B}$ and W_j in the following Equation (15). This equation is derived by taking a linear conditional expectation of both sides of Equation (14), where we are assuming we are taking this linear expectation by conditioning only on $\boldsymbol{X}$ and a matrix $\boldsymbol{F_j}$ of dummies for each LWA $\boldsymbol{j}$, and are omitting $\boldsymbol{S}$.

(15) $E(Y_{ij} \mid \boldsymbol{X}_{ij}, \boldsymbol{F_j}) = \boldsymbol{B}^*\boldsymbol{X}_{ij} + \boldsymbol{W}^*_j\, \boldsymbol{F_j} = (\boldsymbol{B} + \boldsymbol{DG}_x)\boldsymbol{X}_{ij} + (\boldsymbol{W_j} + \boldsymbol{DG}_j)\, \boldsymbol{F}_j$

Equation (15) indicates that the ***B*** and ***W*** coefficients are biased by adding an asterisk to these coefficients. The biased coefficients are equal to the true coefficients plus an additional term, which in part involves the coefficients ***D*** on the omitted variables ***S***.

$\boldsymbol{G}_x$ and $\boldsymbol{G}_j$ are the coefficients that would be estimated by auxiliary regressions of ***S*** on ***X*** and the matrix of fixed effects for each LWA indexed by *j*. Among other things, we know that in these auxiliary regressions it must be true that

(16) $\boldsymbol{G}_j = \text{mean}\,(\boldsymbol{S}_j) - \boldsymbol{G}_x\,(\text{mean}\,(\boldsymbol{X}_j)),$

or the coefficient $\boldsymbol{G}_j$ on LWA *j* in this auxiliary regression will be equal to the mean in LWA *j* of the omitted variables ***S*** minus the regression coefficients from trying to predict ***S*** with ***X***, times the mean in LWA *j* of ***X***.

The point here is that the magnitude of this omitted variable bias depends not only on the size of ***D***, that is on whether the omitted variables ***S*** are important variables in explaining *Y*, but also on $\boldsymbol{G}_j$ being large, which requires that ***S*** vary across LWAs in a manner that cannot be predicted by the variation in ***X*** across LWAs. For example, perhaps individual motivation does affect an individual's success on the common measures, but it is difficult or impossible to measure, and therefore is an important omitted variable ***S*** with a large ***D*** coefficient. But the omitted variable "motivation" will not cause many problems if its variation across LWAs can be well-explained by variables that are included in the adjustment model equation, such as the individual's education and prior employment and earnings history.

Even if there is a large bias in using the adjustment model to estimate value added and performance standard adjustment, the practical issue is whether this bias is larger than is implicit in using the common measures without any adjustment. The use of common measures without adjustment can be seen as estimating value added by estimating Equation (14) but with only the matrix of LWA fixed effects included, and with both the $\boldsymbol{X}$ and $\boldsymbol{S}$ variables omitted from the equation. This equation will consistently estimate the following parameters, derived by taking the conditional expectation of both sides of Equation (14):

(17) $E(Y_{ij} \mid \boldsymbol{F}_j) = \boldsymbol{W}^{**}_j \boldsymbol{F}_j = (\boldsymbol{BH}_j + \boldsymbol{DL}_j + \boldsymbol{W_j}) \boldsymbol{F}_j.$

Here, the $\boldsymbol{H}$ and $\boldsymbol{L}$ coefficients come from auxiliary regressions of $\boldsymbol{X}$ and $\boldsymbol{S}$, respectively, on the matrix of fixed effects for each LWA. If these auxiliary regressions are normalized by setting the weighted sum of the resulting coefficients to zero, then $\boldsymbol{H}_j$ and $\boldsymbol{L}_j$ will be simply the differential of the means of $\boldsymbol{X}$ and $\boldsymbol{S}$ in LWA j from the state means of $\boldsymbol{X}$ and $\boldsymbol{S}$.

Equation (17) looks like it has greater bias in estimating $\boldsymbol{W_j}$ than Equation (15), as it has an additional bias term, $\boldsymbol{BH}_j$. Equation (17) will only lead to less bias if the bias from omitting $\boldsymbol{X}$ goes in the opposite direction from the bias in omitting S. Suppose we are defining variables so that all variables have a positive effect on Y, e.g., both $\boldsymbol{B}$ and $\boldsymbol{D}$ are positive. Then we would only expect the bias from omitting $\boldsymbol{X}$ to offset the bias from omitting $\boldsymbol{S}$ if the omitted variables $\boldsymbol{S}$ are distributed across LWAs in a manner that is negatively correlated with the distribution across LWAs of the included variables $\boldsymbol{X}$. For example, if the unobserved variable "motivation" tends to be higher in LWAs that have lower values of the included variable "prior earnings," then perhaps omitting "prior earnings" in Equation (17) could actually result in improved estimates of value added. In that case, when we add in an adjustment factor for average "prior earnings" in the LWA, we will increase the estimated "value added" for LWAs that have low average "prior earnings." But this adjustment may over-adjust for LWA characteristics if the omitted variable "motivation" tends to be higher in the LWA with low prior earnings. If this over-adjustment is bad enough, we would be better off doing no adjustment at all. On the other hand, if the over-adjustment is not too bad, we would still be better off doing the adjustment.

Although it is theoretically possible for no adjustment to be preferable to adjustment using only observable variables, in the real world we would expect there to be a tendency for the omitted variables that positively affect the common measures to be distributed across individuals, and hence LWAs, in a manner that is positively correlated with included variables that positively affect the common measures. In particular, the ***X*** variables in our adjustment model include the individual's prior employment and earnings, which are essentially lagged values of the labor market common measures. It seems likely that any omitted variable that positively affects the labor market common measures is likely to be positively correlated with the individual's prior employment and earnings. Therefore, it seems more plausible that adjusting for at least some variables will improve our ability to measure the relative value added of different LWAs. Omitted variables will lead to biases in the estimated adjustment, but the biases are likely to be greater with no adjustment at all.

<u>Omission of peer group effects</u>. A criticism of the model that has more force is that the model omits a particular type of variable, the average characteristics of other participants. With the exception of the control for the change in the local unemployment rate, the individual's success on the common measures is assumed to be affected only by the individual's own characteristics. We do not control for the possibility that the individual's own success on the common measure might not only depend on the individual's own values of prior earnings, education, etc., but also on the average value of these variables for other participants in the program in that LWA.

These peer group influences are thought to be particularly important in education. In K–12 education, and even in college education, we can certainly see how any individual's education may depend upon the characteristics of that individual's peers. For example, peers may easily affect the learning climate in a classroom.

But peer group effects are less likely to be important for workforce programs given how these programs are currently structured. The particular services provided to individual program participants are often highly individualized. When education and training is delivered, it is not necessarily delivered with the LWA participants isolated in their own segregated classes. The theoretical case for peer group effects in workforce programs is not strong.

Estimating peer group effects would pose some difficulties. First, given that we are trying to do the estimation using data from only one state, we face the problem of having only a small number of LWAs. Peer group effects would be estimated based on the variation in peer groups across LWAs, so a small number of LWAs means that precise estimates would be difficult to obtain, particularly if we were trying to explore many possible peer group characteristics that might have effects.

Second, it is plausible that average peer group characteristics might be correlated with true value added. It would be no shock to discover that in LWAs with a greater proportion of disadvantaged program participants, local public institutions may tend to be less productive because local voters may put less pressure on local political leaders for high LWA performance. Therefore, controlling for average participant mix may tend to over-adjust for participant characteristics and absorb some true value-added. There are no easy statistical solutions to this problem.

Comparing this adjustment model with alternatives

We briefly consider two alternatives to this paper's adjustment model: the JTPA performance standards system, and negotiated standards based on improvements from past performance.

JTPA performance standards adjustment system. The JTPA performance standards system based adjustments to standards on national regressions that used grouped data from different local areas on each area's performance, participant characteristics, and local economic characteristics. Using grouped data does not fully exploit the information we have on how individual characteristics affect labor market and educational success. For example, suppose there is some individual characteristic that has large effects on how an individual does on one or more common measures, but suppose further that for most LWAs, the LWA mean for that characteristic is close to the overall mean for that characteristic, with only a few LWAs where the LWA mean is significantly different from the overall mean. Using grouped data, it may be difficult to precisely estimate how such a characteristic affects performance on the common measure, and this characteristic may even be dropped from the final specification. This may result in significant biases in measuring value added for the LWAs that happen to have an unusual participant mix for this characteristic.

On the other hand, this grouped data estimation does implicitly reflect peer group effects. The estimated effects of an area's average characteristics on an area's average performance will include both the effects of an individual's characteristics on the individual's performance, and the effects of the characteristics of the individual's peer group on the individual's performance. However, the grouped data estimation does not allow for the individual effects to be distinguished from the peer group effects. Also, as noted above, including peer group effects in some cases mistakenly absorbs a portion of true value-added.

Finally, estimation using mean values for different areas is difficult to do using data from only one state. If we want to have a performance adjustment system that is estimated and matched to the particular characteristics and needs of that state's workforce system, we are almost forced to use individual program participant data in estimating adjustment models.

<u>Negotiated adjustment using past performance</u>. Another alternative that has sometimes been advocated is negotiating performance standards using the LWA's own past performance as a gauge (Baj, undated). This approach may have some advantages under two assumptions: (1) an assumption that the true value-added differences across LWAs are small relative to the large difference in omitted individual characteristics and local economic conditions affecting performance; (2) an assumption that such omitted variables in an LWA do not change much over time.

To show the possible strengths and weaknesses of negotiating standards based on past performance, suppose that the proposed performance standard for an LWA is that the LWA improve by a certain amount over its historical performance, or

(18) $\text{mean}\,(Y_j) - \text{mean}\,(Y_{jb}) \geq K,$

where the j subscript indicates the current period for LWA j, and the jb subscript indicates the historical or "before" period for LWA j.

But suppose that Equation (14) is valid, and we can estimate true value added for each LWA if we include all omitted variables $\boldsymbol{S}$. Then we know that

(19) $\text{mean}\,(Y_j) = \boldsymbol{B}\,\text{mean}\,(\boldsymbol{X}_j) + \boldsymbol{D}\,\text{mean}\,(\boldsymbol{S}_j) + W_j$

(20) $\text{mean}\,(Y_{jb}) = \boldsymbol{B}\,\text{mean}\,(\boldsymbol{X}_{jb}) + \boldsymbol{D}\,\text{mean}\,(\boldsymbol{S}_{jb}) + W_{jb}$

Substituting into Inequality (18) and rearranging, we get the result that standard (18) is equivalent to the following standard for true value added in the current time period:

(21) $W_j \geq K + \boldsymbol{B}(\boldsymbol{X}_{jb} - \boldsymbol{X}_j) + \boldsymbol{D}(\boldsymbol{S}_{jb} - \boldsymbol{S}_j) + W_{jb}$.

Therefore, this standard makes sense if the change in included and omitted $\boldsymbol{X}$ and $\boldsymbol{S}$ is small, and there is little variation across LWAs in value-added during the historical period. On the other hand, if the change in $\boldsymbol{X}$ or $\boldsymbol{S}$ is large, and there was a great deal of variation in value added during the historical period, then basing standards on previous performance would not do a good job of identifying LWAs with the highest true value added in the current period.

It may be possible to control for the change in included variables $\boldsymbol{X}$ to some degree without regressions by matching up the jth LWA with other LWAs that appear to have had similar changes in $\boldsymbol{X}$. If the omitted variables $\boldsymbol{S}$ are correlated with the variables $\boldsymbol{X}$ we observe, such matching may even control to some extent for changes in $\boldsymbol{S}$. However, even if the change in $\boldsymbol{X}$ and $\boldsymbol{S}$ can be controlled for, or happens to be small, we still have the problem that this performance standards system uses an area's own past performance as a yardstick. To the extent to which this past performance reflects true value-added, the result is that LWAs that have previously been able to achieve very high value-added will be subject to a more stringent performance standard than LWAs that previously have had low value-added. This seems unfair. Furthermore, such a procedure seems unlikely to identify exemplary LWA practices or identify the LWAs that most need to be reformed.

Results

To illustrate the adjustment methodology proposed in this paper, models were estimated for various workforce programs in the state of Michigan. Currently, estimates have been completed for four programs (WIA Adult, Employment Service, WIA Dislocated Workers, and WIA Youth), and 14 common measures (four each for the adult groups, two for WIA youth). Estimation is currently underway for two other programs (TANF welfare to work, Trade Adjustment Assistance). Because of the many programs and measures considered, we only present summary results for all the different models and detailed results for a single program common measure, common measure 1 (Entered Employment Rate) for the WIA Adult program.

Table 2 presents sample means for the individuals in the four different programs. The variables reported in Table 2 include all the $\boldsymbol{X}$ variables considered for the adjustment and forecasting regressions, although the actual final regression estimates only retain variables that have a t-statistic whose absolute value is greater than 1.4 and with the expected sign. Some observations based on these means include the following.

- For most of the common measures, the majority of program participants are successful in attaining the program goals of entering a job by one quarter after exit, being retained in some job, and gaining some educational credential.

- The results for the earnings common measures are more mixed. In particular, common measure 4, the earnings gain from one quarter after exit to three quarters after exit for those employed one quarter after exit, tends to be negative. The negative mean for common measure 4 is no surprise, because this measure is by design selecting only program participants who are relatively successful, in the sense of being employed one quarter after program exit. Some of this success turns out to be temporary, an example of what statisticians call "regression to the mean", and the average individual in this selected sample tends to lose earnings by the third quarter after exit. But this negative mean for common measure 4 is a big political problem, because it may not be politically feasible to set a performance standard for earnings gains that is negative. While some measure of post-program earnings changes is useful, policymakers should consider some redefinition so that this measure is unlikely to have a negative mean. For example, this earnings gain measure could be defined only for individuals employed both by the first and third quarters after exit. The extent to which individuals lose jobs between the first quarter and third quarter after exit is already reflected in common measure 2, the job retention rate, so this redefinition would also make common measure 4 more independent of common measure 2.

- As one would expect, the WIA Youth participants are the youngest, followed by WIA Adult, ES, and then WIA Dislocated Workers.

- ES and WIA Displaced Workers tend to have more males, WIA Youth more females, with WIA Adult a more even gender mix.

- All the programs except for ES have a majority of participants who are African-American, whereas the ES participants are majority non-Hispanic white.

- The education of program participants is not quite as low as one might expect, as most program participants have at least a high school diploma, but few participants have a college degree.

- The program participants have quite modest prior quarterly earnings, with WIA Dislocated Workers the highest in prior earnings, followed by ES participants.

- With the exception of WIA Dislocated Workers, the other groups include a sizable proportion of participants with relatively little in prior employment experience.

- Prior to entering the program, WIA Displaced Workers had a large proportion employed in manufacturing, and WIA Youth tended to be employed in food services and retail trade, whereas the other two groups are spread across manufacturing, administration, and retail trade.

- A surprisingly high proportion of program participants were employed at registration.

- During the time period encompassed by this study, the unemployment rate generally increased over a two quarter period by perhaps 1%.

- Most program participants seem to be employed at exit, with hourly wages that are higher than the minimum wage on average but still quite modest. Displaced workers have the highest exit wages, WIA Youth the lowest. Most of those employed at exit are employed close to full-time. The occupations of employment include services and office occupations, but also some participants are in production occupations, with the exception of youth.

Table 3 summarizes the parameter estimates for the "***X***" variables for all 14 adjustment models estimated (three programs with four common measures each, one program with two common measures). Table 3 does this by ranking, for each model, which class of variable (where each class of variable may include only one variable, or several, for example there is only one gender variable but several race variables) has the t-statistic with the greatest absolute value, the second greatest, etc. Table 3 also counts, for each class of variables, the number of models for which at least one variable is statistically significant (t-statistic greater than 1.96 in absolute value). The fixed effects for the LWAs are not considered in any of these rankings.

Over all 14 models, prior employment and prior wages tend to be most often important in explaining the common measures, as judged by the magnitude of t-statistics. Prior employment is particularly important in predicting how individuals will do on common measure 1, whether the individual is employed one quarter after exit. The change in the unemployment rate variable is not always statistically significant, but when it is, it is often an important predictor, particularly in predicting individuals' earnings gains. The race of the individual is seldom statistically significant, and when it is significant, this class of variables is usually not the most important in explaining how individuals do on the common measures. Other classes of variables are of middling significance and importance.

Common measure 4, the individuals' earnings gains from one to three quarters after program exit, is clearly the most difficult common measure to predict. Few of the individual level variables have any statistical significance in explaining this common measure. This suggests that the adjustment approach of this paper will run into some problems for this common measure. For example it will be difficult to do accurate "real-time" predictions of the final adjustment for common measure 4 before the change in unemployment is known.

Table 4 presents the actual parameter estimates and t-statistics for common measure 1, entered employment in first quarter after exit, for the WIA Adult program. The coefficients on the individual explanatory variables are of moderate size, perhaps more moderately sized than would be expected. For example, the highest t-statistic in Table 4 is for the coefficient on the dummy variable for "no wages in all quarters 3 to 12 before registration," yet a "one" on this variable, compared to the omitted dummy of working 6 to 10 of these quarters, is estimated to reduce the probability of being employed one quarter after exit by only 13%. This is an important effect, but it is also perhaps surprisingly low. Similarly, an extra $5,000 in prior average quarterly earnings only increases the probability of being employed one quarter after exit by about 7%, which is important but also perhaps a bit lower than might have been anticipated. Other individual demographic characteristics also have effects that are perhaps a bit lower than one would expect.

However, we would still regard all these estimated effects of individual characteristics on the entered employment rate as plausible. "Any employment" during the first quarter after exit is a fairly minimal program goal that many individuals might achieve, even if they have a seemingly inauspicious background. It must be remembered that these are individuals who have volunteered to participate in this program. Even if these program participants have zero or very low prior earnings and employment experience, the mere fact of their participation indicates a desire to change this situation. Therefore, individual characteristics that might appear to be severely handicapping may not have so severe an effect on achieving the minimal program goal of some employment in the first quarter after exit.

As for the unemployment rate, it is clearly statistically significant, yet it has only a modest "one-to-one" effect: a 1% increase in the local unemployment rate reduces an individual's probability of working one quarter after exit by only about 1%. Local economic conditions do have statistically significant effects on performance on common measure 1, but it would take a huge recession to drive common measure 1 down by a lot. We know from prior studies that more disadvantaged individuals tend to have their employment rates respond more than average to changes in unemployment (e.g., Bartik 2001 and the literature reviewed therein). However, the individuals who participate in the WIA Adult program are a selected group of program participants who are more motivated to achieve increased employment and earnings. In addition, as the unemployment rate increases, the group of program participants may tend to include more individuals who have unobserved characteristics that are correlated with greater labor market success, which will also tend to reduce the measured effects of local unemployment on the entry rate common measure.

Therefore, it is not surprising that even after controlling for these many individual characteristics and the local unemployment rate, there remains considerable variation in common measure 1 performance to be explained by the fixed effects for each LWA. LWA effects range from 12% below to 10% above the state average, even after controlling for all measured demographic differences across LWAs and differences in local unemployment trends.

Table 5 contains further analyses of the results by looking at the predictions and actual values of the common measures for each of the 25 LWAs in Michigan. As implied by Equation (8), the differential performance of each LWA on each common measure can be exactly divided into the sum of two components: (1) the predicted value of the common measure based on that LWA's ***X***s, minus the statewide average of that common measure, where this difference depends on the estimated coefficients on the ***X*** variables times the difference between the ***X*** variables in the LWA and the state; and (2) the value added or ***W*** fixed effects. This first component is what we have referred to as the "adjustment" factor, which the model implies should be subtracted from differential LWA performance to estimate value-added. As can be seen in Table 5, the adjustment factor for each LWA does tend to be positively correlated with the LWA's actual differential performance, particularly for the job entry and job retention common measures. The differential performance of common measure 4, the post-exit earnings gain, is not as well-predicted by each LWA's adjustment factor. However, for each common measure and program considered, the standard deviation of the estimated value-added across the 25 LWAs is always greater than the standard deviation of the adjustment factor across the 25 LWAs.

We interpret these results in Table 5 as follows: Adjustment for differences in the mix of program participants or local economic conditions is important in explaining LWA differential performance. However, a majority of the variation in LWA performance does appear to be due to variations in value-added.

This conclusion is consistent with the visual evidence of Figure 1, which for common measure 1 (job entry) for the WIA Adult program shows the estimated differential performance for each of the 25 LWAs in Michigan, the estimated differential that is due to differences in the adjustment factor for customer mix and the local economy, and the estimated differential due to differences in LWA value-added or productivity. Although more of the variation in LWA performance does reflect estimated value added, there are some LWAs, for example LWA M, in which LWA performance is below the state mean, yet the estimated value added for the LWA is above the state average.

The adjustment factor that has been presented or analyzed so far is the "final adjustment" that is possible after we know the change in local unemployment, and also know who will actually be in the final sample that "counts" for calculating each common measure. As discussed above, LWA managers and state policymakers may find it useful to have an "adjustment factor" for each LWA that can be calculated as individuals enter the program. As individuals enter the program, this estimated adjustment factor will give LWA managers and state policymakers some idea of what types of adjustments will likely be done to raw performance, when the final assessments are done of whether performance standards are met. As outlined before, this can be done by predicting which individuals will be in the final sample for each common measure, and predicting the change in unemployment. For purposes of this paper, we assume that the change in unemployment is predicted to be zero. In the real world, it might be possible to get a better forecast.

Table 6 shows the correlations for each program and common measure between the "final" adjustment factor and this adjustment factor that is predicted at registration. The predictions are generally very good except for common measure 4. Therefore, it seems quite feasible for LWA administrators and state managers to know approximately what the adjustments will be as soon as individuals enter the program.

But LWA administrators will also want to know as soon as possible how they are likely to fare on the performance standard. As outlined above, this can be ascertained by predicting the LWA's values of the common measure using intermediate outcomes, and then using estimated adjustment factors for each LWA to yield an estimated value-added for any program and common measure for each LWA.

As shown by Table 7, intermediate outcomes and other variables often can be used to successfully predict performance on the common measures, and these intermediate outcomes often add significantly to our ability to predict performance on the common measures. This is particularly true for common measure 1. For this common measure, the intermediate outcome of whether the individual is employed at exit is an extremely good predictor of whether the individual is employed one quarter after exit.

These predictions of common measures using intermediate outcomes can be combined with estimates of the adjustment factors to produce an estimate of value added for each LWA, for any program or common measure. Table 8 shows the correlations of this value added that is estimated at exit with the final estimates of "true value-added", which will only be available when all the data on the common measures is available and the change in unemployment is known. As shown in the Table, the exit estimates are highly correlated with the final estimates for common measures 1 and 3 for the WIA Adult program, moderately correlated for common measure 1 for other groups, and moderately correlated for common measure 3 for WIA dislocated workers. Estimating value added at exit is least successful for common measure 4.

Figure 2 shows exit estimates of value added and final estimates of value added for common measure 1 for WIA adults. As shown in the table accompanying the figure, if the performance standard is that value added is positive, that is the performance standard is state mean performance, then the exit estimates of value added for this program and common measure yield correct predictions of whether the LWA met the standard for 21 out of 25 LWAs.

It should be kept in mind that these exit estimates of value added are available before the change in unemployment is known, and before we have any direct data on the common measure. These exit estimates would be available for all program participants of a particular program year probably at least three quarters of a year before the final estimates of value-added are available, and even longer for some common measures. Furthermore, exit estimates of value-added for a portion of program participants can be calculated throughout the program year as participants exit.

Implementation Issues and Improvements Needed

Implementation in other states requires that similar models be estimated. This should not require anything beyond standard statistical software and econometric analysis.

A state's implementation of this model requires that its wage record data be capable of being integrated with the state's workforce program administrative data on a real-time basis. Prior earnings and employment are clearly very important in predicting many of the common measures. Therefore, these prior earnings and employment data are needed to make the proper adjustments for individual characteristics that are needed to accurately estimate program value added. Data on prior earnings and employment are best gathered on a consistent basis from wage record data.

Once estimated coefficients are available and a real-time data base is created that integrates wage record and administrative data, the model could be placed on a desktop and largely automated. It could be run simply as a "black box," allowing program managers to at any point click on a few icons on a computer and get updated data on estimated adjustment factors and estimated value added for the current program year. Alternatively, the model could be put into a spreadsheet, just as JTPA was, to provide more transparency for how different factors contribute to the model results.

The most needed improvements to this model are better predictions of the common measures using intermediate outcomes. Employment at exit is a good predictor for common measure 1, but we need additional intermediate outcome variables that will better predict the job retention and earnings gains common measures.

One possible predictor of the common measures would be variables measuring the individual's performance in training activities: attendance, "grades," etc. Such data might also be useful to program managers in monitoring and improving training activities. Another possible predictor of the common measures is very short-term follow-up survey data after exit, even only 30 or 60 days, to see whether the jobs held at exit have been retained.

We may also be able to make improvements in adjusting models by better measuring the individual characteristics of program participants as of registration. For example, various psychological screening tests might identify hidden strengths and weaknesses of program participants. In addition to helping improve the accuracy of adjustment models, this additional information at registration might be useful in assigning program participants to services.

As a general principle, it seems likely that many improvements in gathering data that would help predict how program participants would do without the program or that would provide intermediate predictions for how the program affects outcomes, are likely to be useful to program administrators for many reasons, not just the adjustment models outlined in this paper.

Conclusion

The theory and illustrative results in this paper suggest that even with a single state's data, we can adjust for customer mix and local economic conditions to produce credible "value added" measures for LWAs. The estimated effects of individual characteristics and local economic conditions on outcomes are of modest but plausible magnitude. Adjustment for customer mix and local economic conditions does make a difference to measuring value-added, which is a necessary condition for pursuing such an approach. Furthermore, major variations in estimated value-added across LWAs still exist after these adjustments, which is another necessary condition for pursuing this approach.

We can use intermediate outcomes to predict common measures on a real-time basis for some common measures and programs, but not others. Some improvements in data on intermediate outcomes are needed.

Data systems for workforce programs have not been set up with the data needs of performance adjustment models as an important consideration. Improvements in data systems to meet the needs of performance adjustment models might well be broadly useful for many program management purposes.

References

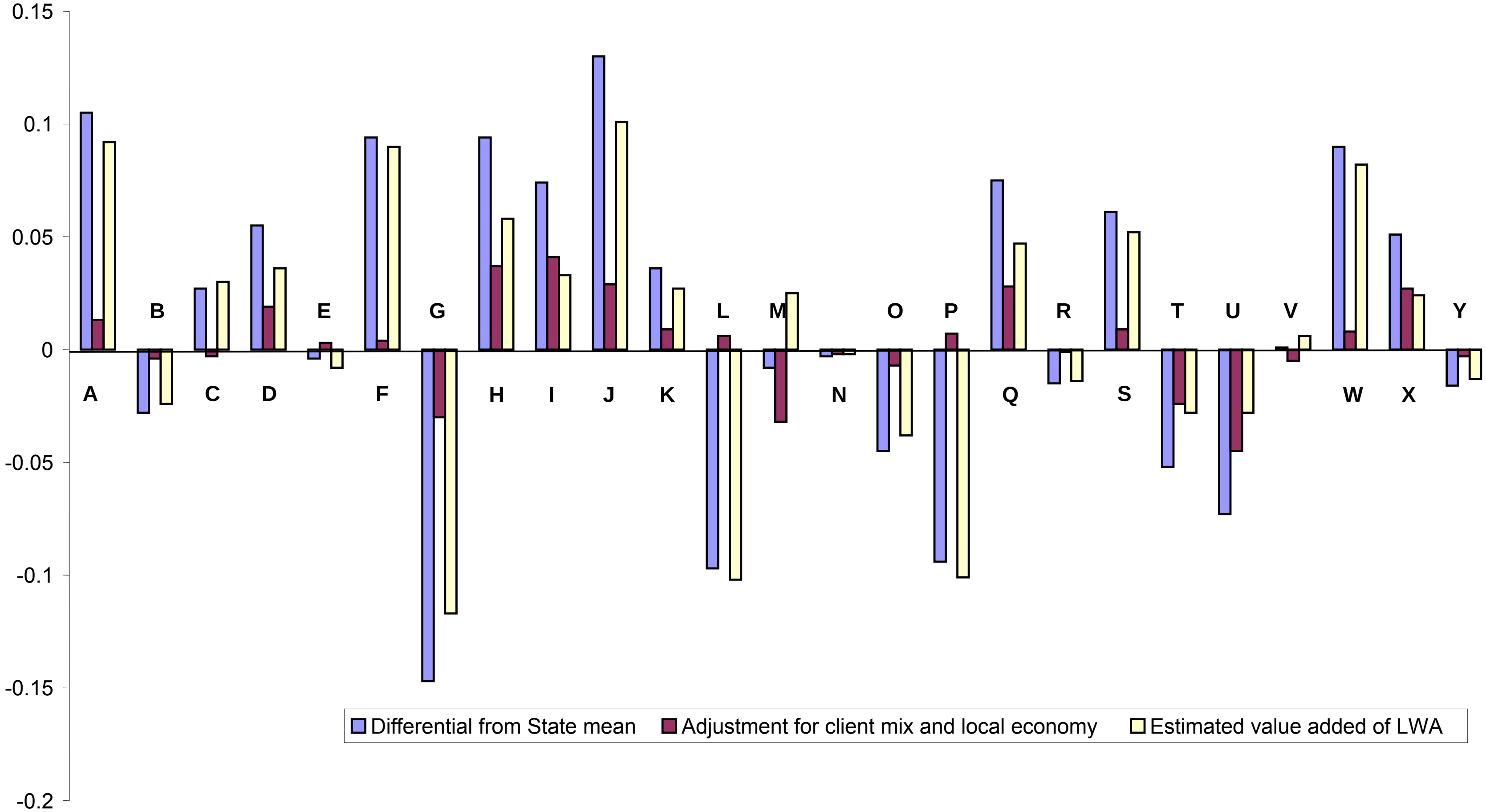
Amemiya, Takeshi. 1985. *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Baj, John. Undated. “Handouts for Adjustment Models Discussion.” Center for Governmental Studies, Northern Illinois University.

Bartik, Timothy J. 2001. *Jobs for the Poor: Can Labor Demand Policies Help?* New York and Kalamazoo, MI: Russell Sage Foundation and W.E. Upjohn Institute for Employment Research.

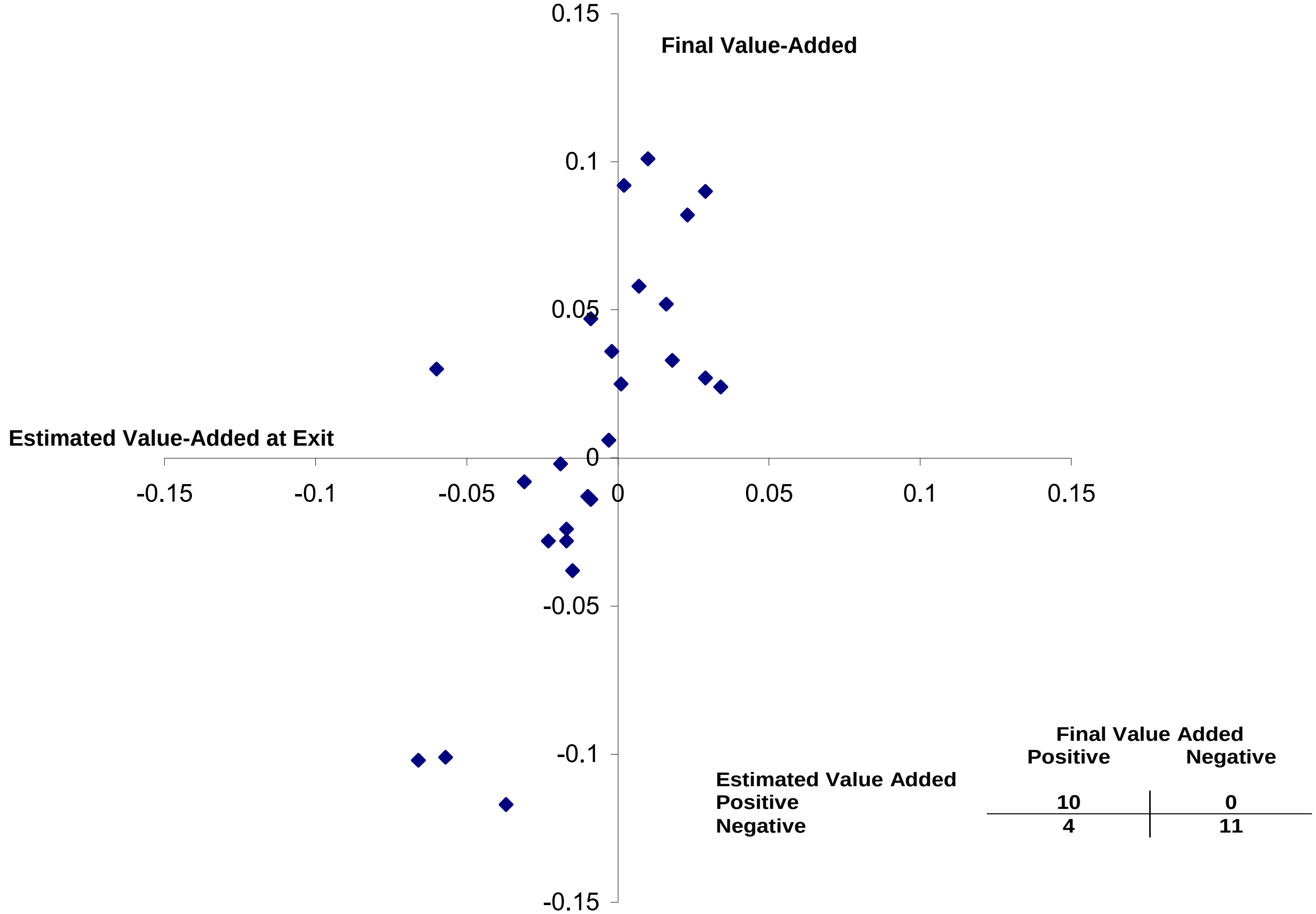
U.S. Department of Labor. 2003. “Common Measures Policy.” *Training and Employment Guidance Letter No. 15-03*. Washington, DC: U.S. Department of Labor, Employment and Training Administration.

Figure 1. LWA Differentials from State Mean, LWA Adjustments, and LWA Value-Added for WIA Adult Program, Common Measure 1 (Job Entry)



NOTE: Each letter and three bars shows results for one of 25 LWAs in Michigan. Mean of "job entry" common measure for WIA Adults is 0.763. To illustrate meaning of chart, LWA A is 0.105 above state mean (0.763 + 0.105 = 0.868), and 0.013 of this differential is explained by client mix and the local economy, 0.092 by "value added."

Figure 2. "Final" Value-Added vs. Estimated Value-Added at Exit for WIA Adult Program, "Job Entry" Common Measure



NOTE: Each diamond represents results for one LWA in Michigan. "Final" Value-Added is estimated value-added after common measure value is known, change in unemployment is known, and final sample is known. Estimated value-added at exit uses intermediate outcomes to predict common measure, assumes no change in unemployment in doing adjustments, and uses probability weights of being in final sample.

Table 1. Brief Definitions of Common Measures for U.S. Workforce Programs, Including Both Measures for Which This Paper Estimates Adjustment Models, and Measures Not Analyzed By This Paper

Name and label of common measure	Brief Definition	Adjustment model estimated in this study, and for what groups?
Adult common measures		
Common measure 1: Entered employment	Of those not employed at registration in program, the proportion employed in the first quarter after exit from the program, based on wage record data.	Yes: WIA Adult, Employment Service (ES), WIA Dislocated Workers, TANF*, TAA*
Common measure 2: Job retention	Of those employed in the first quarter after exit from the program, the proportion employed in both the second and third quarters after exit.	Yes: WIA Adult, ES, WIA Dislocated Workers, TANF*, TAA*
Common measure 3: Pre to Post Earnings Change	Of those employed in the first quarter after exit from the program, the percentage earnings gain from the first quarter before registration to the first quarter after exit.	Yes: WIA Adult, ES, WIA Dislocated Workers, TANF*, TAA*
Common measure 4: Post earnings change	Of those employed in the first quarter after exit from the program, the percentage earnings gain from the first quarter after exit to third quarter after exit.	Yes: WIA Adult, ES, WIA Dislocated Workers, TANF*, TAA*
Youth common measures		
Common measure 1: Entered employment or advanced education/training	Of those in secondary school at registration, and those not in secondary school who are also not in post-secondary education, employment, or military, the proportion who during first quarter after exit are either employed, or enrolled in post-secondary education or advanced training, or in military. Persons in secondary school at exit are excluded.	Yes: WIA Youth
Common measure 2: Attainment of educational/ training credential	Of those in education or technical/occupational training at registration, or during program, the proportion who attain a diploma, GED, or certificate by the end of the 3rd quarter after exit. Persons in secondary school at exit are excluded.	Yes: WIA Youth
Common measure 3: Literacy or numeracy gains	Of those who are basic skills deficient when pre-tested, and who either are in program for year or exit from program, the proportion who advance at least one education functioning level in any skill area (reading, writing, numeracy, speaking, listening, functional, workplace skills).	No: These data are not available yet in Michigan
Adult and youth common measures		
Efficiency measure	Spending divided by program participants	No: unclear whether adjustment is feasible

Note: For all common measures, program participants are excluded if at exit, or during three quarters after exit, the participant is in prison or hospital, providing care to family, deceased, or a reservist called to active duty.
*TANF and TAA analyses not yet completed.

Table 2. Sample Means for Four Michigan Workforce Programs

Variable	Adult WIA	ES	WIA dislocated workers	Youth WIA
Sample size, common measure 1	10,274	87,389	7,599	3,248
Sample size, Other common measures	9,056	16,946 (50,710 for CM3)	6,284	1,973
Common measure:				
1. Employed 1 quarter after exit (of those not employed at registration)	0.763	0.504	0.801	
2. Retained job in quarters 2 and 3 after exit	0.726	0.735	0.839	
3. Percentage earnings change from quarter before registration to 1 quarter after exit	102.5	−2.1	23	
4. Percentage earnings change from 1 quarter after exit to 3rd quarter after exit (Measures 2, 3 and 4 only include those employed 1 quarter after exit. Percentage earnings change for individual is change in earnings divided by state mean in base period.)	−13.4	−0.6	−6.1	
Youth common measures:				
1. Employed one quarter after exit, or exited due to entering military, apprenticeship, training or post-secondary education. Excluded if employed and not in secondary education at registration.				0.656
2. Of students at registration, or received training/education during program, whether attained diploma or other education/training credential				0.657
Age				
29 or less	0.370	0.280	0.165	
30–39	0.305	0.277	0.283	
40–49	0.226	0.247	0.348	
50 or more	0.099	0.206	0.204	
Age at registration = 14				0.015
Age at registration = 15				0.023
Age at registration = 16				0.063
Age at registration = 17				0.151
Age at registration = 18				0.175
Age at registration = 19				0.233
Age at registration = 20				0.191
Age at registration = 21				0.148
Gender				
Male	0.49	0.61	0.55	0.41
Female	0.51	0.39	0.45	0.59

Table 2. (Continued)

Variable	Adult WIA	ES	WIA dislocated workers	Youth WIA
Race				
White	0.194	0.670	0.223	0.262
African American	0.802	0.237	0.773	0.735
Hispanic/Latino	0.020	0.068	0.011	0.029
Native American/Alaskan Native	0.034	0.022	0.020	0.042
Other (Asian/Hawaiian/Pac. Islander)	0.007	0.012	0.008	0.010
Education				
Less than high school	0.154	0.141	0.063	0.634
Certificate equivalent to HS	0.119	0.080	0.094	0.044
High school graduate/GED	0.509	0.351	0.554	0.305
Some college	0.166	0.306	0.192	0.018
Bachelor degree	0.045	0.092	0.078	0
Advanced	0.007	0.031	0.019	0
Wages				
Avg. quarterly wages in non-zero quarters 3–12 before registration (in thousands)	2.856	5.482	6.269	0.792
Wages zero all 10 quarters (3–12 before registration)	0.139	0.124	0.071	0.212
1–5 non-zero wage quarters	0.241	0.158	0.095	0.460
6–10 non-zero wage quarters	0.620	0.718	0.834	0.328
Has disability	0.070	0.015	0.025	0.140
Veteran	0.066	0.184	0.106	0.003
Single parent	0.306		0.144	0.296
Long-term TANF	0.180		0.016	0.205
General/refugee/SSI assistance	0.055		0.011	0.085
Food-stamp recipient	0.305		0.056	0.291
Homeless	0.020		0.004	0.026
Pregnant or parenting youth	0.007		0.001	0.320
Limited English	0.026		0.038	0.016
Displaced homemaker	0.003		0.088	0
Offender	0.052		0.008	0.123
Other barriers to employment	0.034		0.013	0.160
Number in family	2.2		2.3	2.2
Alternate or no phone only	0.037	0.044	0.013	0.04
Not registered for selective service	0.016		0.03	0.01
Layoff/termination		0.551		
Plant closure		0.034		
Long-term unemployed		0.011		
Self-employed, farmer		0.001		

Table 2. (Continued)

Variable	Adult WIA	ES	WIA dislocated workers	Youth WIA
Basic skills deficiency				0.617
Behind 1 grade level				0.348
Prior industry				
Agriculture, forestry, fishing	0.006	0.012	0.003	0.006
Mining	0.001	0.002	0.001	0
Utilities	0.001	0.002	0.001	0
Construction	0.029	0.082	0.022	0.013
Manufacturing	0.190	0.180	0.437	0.045
Wholesale trade	0.026	0.039	0.046	0.010
Retail trade	0.125	0.100	0.079	0.161
Transportation, warehousing	0.015	0.028	0.025	0.006
Information	0.007	0.017	0.011	0.005
Finance and insurance	0.013	0.022	0.024	0.003
Real Estate, rental, leasing	0.010	0.014	0.007	0.008
Professional, scientific, technical	0.027	0.055	0.047	0.009
Company/enterprise mgt	0.003	0.003	0.001	0.003
Admin, support and waste mgt	0.175	0.124	0.103	0.110
Educational services	0.020	0.020	0.014	0.049
Health care/social assistance	0.068	0.051	0.035	0.039
Art, entertainment, recreation	0.011	0.014	0.008	0.019
Accommodation and food services	0.090	0.061	0.024	0.266
Other services (except public admin)	0.021	0.024	0.016	0.016
Public administration	0.011	0.012	0.010	0.012
Unclassifiable	0.003	0.003	0.004	0.004
Industry missing	0.009	0.009	0.010	0.003
Employed at registration (only relevant for CMs 2 through 4)	0.130	0.151	0.031	0.139
Change unemployment rate, (registration − 1) quarter to (exit + 1) quarter	0.009	0.007	0.009	0.008
Change unemployment rate, exit + 1 quarter to exit + 3 quarters	0.007	0.012	0.007	
Variables used in exit models only:				
Employed at exit	0.881		0.897	0.638
Hourly wage at exit	9.25		11.40	7.32
Weekly hours	37.3		38.9	33.9

Table 2. (Continued)

Variable	Adult WIA	ES	WIA dislocated workers	Youth WIA
Exit occupation:				
Management, business, financial	0.028		0.058	0.010
Professional and related	0.048		0.056	0.021
Services	0.179		0.086	0.196
Sales and related	0.053		0.046	0.077
Office and administrative support	0.110		0.134	0.093
Farming, fishing and forestry	0.001		0.001	0.004
Construction and extraction	0.020		0.023	0.012
Installation, maintenance and repair	0.019		0.038	0.009
Production	0.169		0.238	0.069
Transportation and material moving	0.078		0.095	0.023
Missing or military	0.176		0.123	0.122
ES service means as of exit:				
Resume assistance/preparation		0.260		
Specific LMI		0.303		
Veterans vocational guidance		0.031		
Provided case management		0.001		
Referral, supportive service		0.115		
Other testing		0.007		
Referred to training		0.007		
Enrolled in training		0.002		
Job development		0.037		
Job search planning		0.109		
Job search workshop		0.023		
Referred to WIA services		0.084		
Job referral		0.060		

NOTE: For three WIA groups, the participant must have registered and exited between July 1, 2000 and September 30, 2002. For ES, the individual must have exited between July 1, 2002 and March 31, 2003 for common measures 1 and 3, and between July 1, 2002 and September 30, 2002 for common measures 2 and 4. The sample means reported above are generally for the sample used to estimate common measure 1 (sample means for same program and other common measures are similar).

Table 3. Summary of Statistical Significance and Relative Importance of Different Classes of Variables for 14 Adjustment Models

	Common measure 1 Entered employment (4 models total)	Common measure 2 Job retention (3 models total)	Common measure 3 Pre- to post-earnings change (3 models total)	Common measure 4 Post-earnings change (3 models total)	Youth common measure 2 Gained educational credential (1 model total)	Overall summary: Number of models in which significant (out of 14); average ranking across all models in which significant
Gender	3, 9, 4, __	2, 8, 2	3, 5, 6	__, __, __	3	10; 4.5
Race	__, 8, 7, __	__, 7, __	__, __, 5	__, __, __	4	5; 6.2
Age	5, 2, 2, 4	__, 4, __	__, 7, __	__, 3, __	2	8; 3.6
Education	7, 4, __, 5	3, 2, 5	5, 8, __	__, 2, __	1	10; 4.2
Prior employment	1, 1, 1, 1	4, 1, __	1, 2, 3	__, 4, __	__	10; 1.9
Prior wages	2, 3, 6, 2	1, 6, 3	4, 1, 1	__, __, 2	__	11; 2.8
Barriers	6, 5, 3, 3	6, 3, 1	__, 4, __	__, __, __	5	9; 4.0
Prior industry	__, 7, 5, __	7, 5, 4	6, 6, 4	2, 1, 3	__	11; 4.5
Change in unemployment	4, 6, __, __	5, __, __	2, 3, 2	1, __, 1	__	8; 3.0

NOTE: For each common measure and class variable, that cell lists ranking/significance results in the following order: WIA Adult, ES, WIA Dislocated, and Youth. For common measures 2 through 4, no youth model is relevant. Obviously, the Youth common measure 2 results are only for that one program. If for a given class of variables, no variable is statistically significant, that class of variables is unranked for that model, which is indicated by an underscore. To determine ranking, we first examine which class of variable has greatest *t*-statistic (in absolute value) for that model, and that class gets rank of one. We then look within that model at other classes of variables, and the class which includes the next highest *t*-statistic (ignoring variables in the class which has already been ranked) is ranked second. The ranking continues along the same logic until all remaining classes have no variables that are statistically significant for that model. For example, the "3, 9, 4, __" in the cell for gender for CM1 means that the gender class of variables is the 3rd most important for the WIA Adult program, 9th most important for ES, 4th most important for the WIA Dislocated program, and insignificant for WIA Youth.

Table 4. Adult WIA Parameter Estimates (*t*-statistics in parentheses)

	Common Measure 1: Job entry	
	Parameter estimate	***t*-statistics**
Dependent variable mean	0.763	
Age		
29 or less	0.060	(3.91)
30–49	0.030	(2.08)
Gender		
Male	−0.045	−(5.16)
Education		
Less than high school	−0.032	(−2.75)
Wages		
Avg. quarterly wages in non-zero quarters 3–12 before registration (in thousands)	0.013	(6.36)
Wages zero all quarters, 3–12 quarters before registration	−0.134	(−9.03)
1–5 non-zero wage quarters	−0.087	(−8.23)
Has disability	−0.061	(−3.59)
General/refugee/SSI assistance	−0.055	(−3.02)
Homeless	−0.069	(−2.33)
Alternate or no phone only	−0.068	(−3.10)
Prior industry		
Construction	−0.046	(−1.87)
Educational services	0.054	(1.85)
Health care/social assistance	0.026	(1.55)

	Common Measure 1: Job entry	
	Parameter estimate	***t*-statistics**
LWA		
A	0.092	(3.10)
B	−0.024	(−1.43)
C	0.030	(0.72)
D	0.036	(2.71)
E	−0.008	(−0.14)
F	0.090	(1.00)
G	−0.117	(−3.98)
H	0.058	(2.69)
I	0.033	(1.06)
J	0.101	(1.99)
K	0.027	(2.28)
L	−0.102	(−3.46)
M	0.025	(1.65)
N	−0.002	(−0.09)
O	−0.038	(−1.27)
P	−0.101	(−5.59)
Q	0.047	(1.32)
R	−0.014	(−1.65)
S	0.052	(1.59)
T	−0.028	(−2.18)
U	−0.028	(−0.87)
V	0.006	(0.28)
W	0.082	(2.87)
X	0.024	(1.00)
Y	−0.013	(−0.42)

	Common Measure 1: Job entry				Common Measure 1: Job entry	
	Parameter estimate	***t*-statistics**			**Parameter estimate**	***t*-statistics**
Change in unemployment rate	−1.018	(−4.14)	(−4.14)			

Table 5. Decomposing LWA Performance into Adjustments for LWA Characteristics and Value-Added

Program and common measure	State Mean	Correlation of predicted adjustment with differential LWA performance	Standard deviation of: LWA mean	Standard deviation of: LWA adjustment factor	Standard deviation of: Estimated LWA value added
WIA Adult					
cm 1: Job entry	0.763	0.702	0.072	0.021	0.059
cm2: Job retention	0.726	0.779	0.064	0.028	0.045
cm 3: Pre to post earnings gain	102.5	0.508	20.2	9.4	17.4
cm 4: post earnings gain	−13.4	0.395	5.4	1.9	5.0
ES					
cm 1: Job entry	0.504	0.530	0.041	0.019	0.035
cm2: Job retention	0.735	0.773	0.053	0.027	0.037
cm 3: Pre to post earnings gain	−2.1	0.532	10.2	7.4	8.9
cm 4: post earnings gain	−0.6	0.162	5.5	3.8	6.2
WIA dislocated workers					
cm 1: Job entry	0.801	0.500	0.069	0.033	0.059
cm2: Job retention	0.839	0.310	0.053	0.016	0.051
cm 3: Pre to post earnings gain	23.0	0.020	15.3	7.1	16.8
cm 4: post earnings gain	−6.1	0.294	3.7	1.3	3.5
WIA Youth					
cm 1: Job entry	0.656	0.488	0.088	0.051	0.077
cm2: Obtain ed credential	0.657	0.397	0.193	0.075	0.177

NOTE: Correlations and standard deviations for each cell are calculated based on 25 observations, one for each LWA. The correlations are based on a variation of Equation (8): (mean Y_j − mean Y_s) = B (mean X_j − mean X_s) + W_j. The correlation is between the left hand side of this equation and the first term, the "adjustment factor." The standard deviations are for the left hand side of the equation, the first expression on the right hand side, and W_j. Because the left hand side and the adjustment factor both subtract out the state mean from the value for each LWA, the correlations and standard deviations involving these terms would also apply if these expressions were replaced by mean Y_j and B (mean X_j).

Table 6. Correlation of "Final" Performance Adjustment with Adjustment Estimate at Registration

	CM1 (Job entry)	CM2 (Job retention)	CM3 (Pre- to post-Earnings gain)	CM4 (Post-earnings gain)	Youth CM2 (Obtain educational credential)
Adult WIA	0.948	0.973	0.903	0.676	
ES	0.940	0.993	0.953	0.866	
Dislocated WIA	1.000	0.993	0.914	0.776	
Youth WIA	0.898				0.922

NOTE: "Final" performance adjustment is $\boldsymbol{B}$(mean $\boldsymbol{X}_j$ − mean $\boldsymbol{X}_s$). This is calculated after sample used for that common measure is known and change in unemployment is known. Estimated performance adjustment at registration uses mean of $\boldsymbol{X}$s for LWA j except that change in unemployment is assumed to be zero. In addition, mean of $\boldsymbol{X}_j$ is calculated as weighted mean of registration sample. Weights used are estimated probabilities from logit estimates of probability of each observation in registration sample being in final sample for that common measure. Correlations use 25 observations, one for each LWA.

Table 7. Correlation of Exit Predictions of Common Measure for LWA with Actual LWA Mean for Common Measure, Compared to Correlation of Registration Prediction with Actual LWA Mean

	CM1 (Job entry)		CM2 (Job retention)		CM3 (Pre- to Post-earnings change)		CM4 (Post-earnings change)		Youth CM2 (Got educational credential)	
	Exit with Actual	Regis. with Actual	Exit with Actual	Regis. with Actual	Exit with Actual	Regis. with Actual	Exit with Actual	Regis. with Actual	Exit with Actual	Regis. with Actual
Adult WIA	0.826	0.595	0.760	0.745	0.717	0.484	0.138	0.133		
ES	0.669	0.565	0.807	0.797	0.562	0.520	0.220	0.250		
Dislocated WIA	0.603	0.500	0.371	0.316	0.195	−0.067	−0.026	0.105		
Youth WIA	0.573	0.299							N/A	0.350

NOTE: Correlations are based on 25 observations, one for each LWA. First set of predictions use “intermediate outcomes,” observed at exit, to predict common measures. Individual predictions are weighted by logit probabilities, estimated at exit, for being in that common measure sample, and weighted means for each LWA are calculated. Correlation is between that weighted mean prediction and actual LWA mean. Second set of correlations are based on similar predictions and logit weights, but estimated at registration.

Table 8. Correlations of Value-Added Estimates at Exit with Final Value-Added Estimates

	CM1 (Job entry)	CM2 (Job retention)	CM3 (Pre- to post-earnings change)	CM4 (Post-earnings change)	Youth CM2 (Got educational credential)
Adult WIA	0.730	0.410	0.627	0.154	
ES	0.382	0.096	0.170	0.004	
Dislocated WIA	0.344	0.218	0.459	−0.079	
Youth WIA	0.316				NA

NOTE: Correlations are based on 25 observations, one for each LWA. The value-added estimated at exit is calculated by adding adjustment estimated at exit to prediction of common measure using intermediate outcomes estimated at exit. The adjustment estimated at exit uses original coefficients, but the weighted means use weights that are based in part on intermediate outcomes. The final value-added estimates are the ex-post estimates, when common measures, final sample, and change in unemployment are known.